

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:)
)
 Toshikazu Hori, et al.)
) Group Art Unit: Not Assigned
 Serial No.:)
) Examiner: Not Assigned
 Filed: September 13, 2000)

JC862 U.S. PTO
 09/661428
 09/13/00



For: CHARACTER RECOGNITION DEVICE AND METHOD FOR DETECTING
 ERRONEOUSLY READ CHARACTERS, AND COMPUTER READABLE
 MEDIUM TO IMPLEMENT CHARACTER RECOGNITION

SUBMISSION OF CERTIFIED COPY OF PRIOR FOREIGN
APPLICATION IN ACCORDANCE
WITH THE REQUIREMENTS OF 37 C.F.R. §1.55

*Honorable Commissioner of
 Patents and Trademarks
 Washington, D.C. 20231*

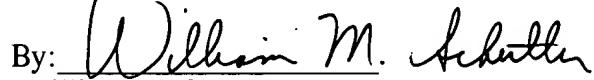
Sir:

In accordance with the provisions of 37 C.F.R. §1.55, the applicant(s) submit(s) herewith a certified copy of the following foreign application:

Japanese Patent Application No. 11-261968
 Filed: September 16, 1999

It is respectfully requested that the applicants be given the benefit of the foreign filing date as evidenced by the certified papers attached hereto, in accordance with the requirements of 35 U.S.C. §119.

Respectfully submitted,
 STAAS & HALSEY LLP

By: 
 William M. Schertler
 Registration No. 35,348

700 11th Street, N.W., Ste. 500
 Washington, D.C. 20001
 (202) 434-1500
 Date: 9/13/2000

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

JC862 U.S. PTO
09/661428
09/13/00

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出願年月日
Date of Application:

1999年 9月16日

出願番号
Application Number:

平成11年特許願第261968号

出願人
Applicant(s):

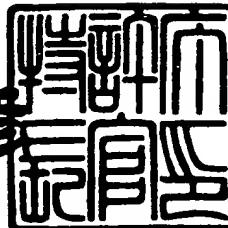
富士通株式会社
ジャンプ セブ インコーポレーション

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年 3月24日

特許庁長官
Commissioner,
Patent Office

近藤 隆彦



【書類名】 特許願

【整理番号】 9951084

【提出日】 平成11年 9月16日

【あて先】 特許庁長官殿

【国際特許分類】 G06K 9/00

【発明の名称】 文字認識装置及びその方法とプログラム記録媒体

【請求項の数】 10

【発明者】

【住所又は居所】 東京都大田区新蒲田一丁目17番25号 株式会社富士
通ラーニングメディア内

【氏名】 堀 外志数

【発明者】

【住所又は居所】 東京都大田区新蒲田一丁目17番25号 株式会社富士
通ラーニングメディア内

【氏名】 吉本 浩二

【発明者】

【住所又は居所】 東京都大田区新蒲田一丁目17番25号 株式会社富士
通ラーニングメディア内

【氏名】 松下 勤

【発明者】

【住所又は居所】 東京都台東区蔵前3丁目1番9号

【氏名】 椎谷 範一

【特許出願人】

【識別番号】 000005223

【氏名又は名称】 富士通株式会社

【特許出願人】

【住所又は居所】 フィリピン国、セブ市 6000、バニラッド、パセオ
アナベル マリア ルイサ エステート パーク、18

【氏名又は名称】 ジャンプ セブ インコーポレーション

【国籍】 フィリピン

【代理人】

【識別番号】 100095072

【弁理士】

【氏名又は名称】 岡田 光由

【電話番号】 03-3807-1818

【選任した代理人】

【識別番号】 100074848

【弁理士】

【氏名又は名称】 森田 寛

【手数料の表示】

【予納台帳番号】 012944

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9707817

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 文字認識装置及びその方法とプログラム記録媒体

【特許請求の範囲】

【請求項1】 イメージスキャナの読み取る文書画像を処理対象として、該文書画像の持つ文字を認識する文字認識装置において、

別々の認識手法に従って、上記文書画像の持つ文字を認識する複数の認識手段と、

上記認識手段の認識結果の不一致個所を抽出する抽出手段と、

上記抽出手段の抽出する不一致個所を明示しつつ、上記文書画像の持つ文字の認識結果を出力する出力手段とを備えることを、

特徴とする文字認識装置。

【請求項2】 イメージスキャナの読み取る文書画像を処理対象として、該文書画像の持つ文字を認識する文字認識装置において、

規定の認識手法に従って、上記文書画像の持つ文字を認識する第1の認識手段と、

上記第1の認識手段の用いる認識手法とは別の認識手法に従って、上記文書画像の持つ文字を認識する第2の認識手段と、

上記第1の認識手段の認識結果と上記第2の認識手段の認識結果との不一致個所を抽出する抽出手段と、

上記抽出手段の抽出する不一致個所を明示しつつ、上記文書画像の持つ文字の認識結果を出力する出力手段とを備えることを、

特徴とする文字認識装置。

【請求項3】 請求項1又は2記載の文字認識装置において、

出力手段は、文書画像と認識結果とを対比させつつ該認識結果を出力することを、

特徴とする文字認識装置。

【請求項4】 請求項1又は2記載の文字認識装置において、

出力手段は、文書画像と認識結果とを対比させつつ該認識結果をディスプレイ画面に表示することで出力するとともに、その表示にあたって、認識結果の表示

領域にカーソルを表示するときに、そのカーソル位置に対応付けられる文書画像の個所を明示する形態で該文書画像を表示することを、

特徴とする文字認識装置。

【請求項5】 請求項1～4のいずれかに記載される文字認識装置において

出力手段は、不一致個所については、認識文字に代えて、その旨を表示する記号又は空白を出力することを、

特徴とする文字認識装置。

【請求項6】 請求項1～4のいずれかに記載される文字認識装置において

出力手段は、認識文字数が等しい不一致個所については、最も評価値の高い認識文字を、一致個所の出力形態とは異なる出力形態で出力することを、

特徴とする文字認識装置。

【請求項7】 請求項1～4のいずれかに記載される文字認識装置において

出力手段は、認識文字数が異なる不一致個所については、規定の基準で選択される該不一致個所の認識文字を、一致個所の出力形態とは異なる出力形態で出力することを、

特徴とする文字認識装置。

【請求項8】 請求項1～7のいずれかに記載される文字認識装置において

出力手段は、認識結果が一致するものの、その認識の信頼性が低い個所については、その旨が分かる形態で出力することを、

特徴とする文字認識装置。

【請求項9】 イメージスキャナの読み取る文書画像を処理対象として、該文書画像の持つ文字を認識する文字認識方法において、

規定の認識手法に従って、上記文書画像の持つ文字を認識する第1の処理過程と、

第1の処理過程で用いる認識手法とは別の認識手法に従って、上記文書画像の

持つ文字を認識する第2の処理過程と、

第1の処理過程の認識結果と第2の処理過程の認識結果との不一致個所を抽出する第3の処理過程と、

第3の処理過程で抽出した不一致個所を明示しつつ、上記文書画像の持つ文字の認識結果を出力する第4の処理過程とを備えることを、

特徴とする文字認識方法。

【請求項10】 イメージスキャナの読み取る文書画像を処理対象として、該文書画像の持つ文字を認識する文字認識装置の実現に用いられるプログラムが格納されるプログラム記録媒体であって、

規定の認識手法に従って、上記文書画像の持つ文字を認識する第1の認識処理と、

上記第1の認識処理の用いる認識手法とは別の認識手法に従って、上記文書画像の持つ文字を認識する第2の認識処理と、

上記第1の認識処理の認識結果と上記第2の認識処理の認識結果との不一致個所を抽出する抽出処理と、

上記抽出処理の抽出する不一致個所を明示しつつ、上記文書画像の持つ文字の認識結果を出力する出力処理とをコンピュータに実行させるプログラムが格納されることを、

特徴とするプログラム記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、イメージスキャナの読み取る文書画像を処理対象として、その文書画像の持つ文字を認識する文字認識装置及びその方法と、その文字認識装置の実現に用いられるプログラムが格納されるプログラム記録媒体とに関し、特に、ユーザが認識結果の誤読文字を効率的にチェックできるようにする文字認識装置及びその方法と、その文字認識装置の実現に用いられるプログラムが格納されるプログラム記録媒体とに関する。

【0002】

近年、オフィスにおけるワークフローの効率化を図るために、文書を電子的にファイリングしていくことが求められている。このようなことを背景にして、電子化されていない文書をイメージスキャナで読み取って、それを文字認識し、誤読文字については手入力で修正していくことで、文書を電子化していくことが行われている。

【0003】

しかるに、現在の技術では、完璧な文字認識を実現する文字認識装置は提供されていないのが実情であり、これがために、このような作業を行う際に必要となる誤読文字の検出に多大な負荷を強いられるという問題点がある。これから、この問題点の解決を実現する技術の構築が呼ばれている。

【0004】

【従来の技術】

電子化されていない文書をイメージスキャナで読み取って、それを文字認識し、誤読文字については手入力で修正していくことで文書を電子化していく場合、ユーザが、電子化対象の文書と文字認識装置により認識されたその文書の認識結果とを比較することで、誤読文字を検出していくことになる。

【0005】

従来では、このような作業を効率的に実行できるようにするための支援技術については、全く提供されていないというのが実情である。

【0006】

これから、従来では、電子化されていない文書をイメージスキャナで読み取って、それを文字認識し、誤読文字については手入力で修正していくことで文書を電子化していく場合には、できるだけ認識率の高い文字認識装置を用いるという方法を探っている。

【0007】

このような高い認識率を実現する文字認識装置の1つとして、特開平7-114620号に開示されるものがある。

【0008】

この特開平7-114620号に開示される文字認識装置では、スタイラスペン等に

より入力される手書き文字を認識する場合に、異なる認識手法により構成された複数の文字認識プログラムを用意して、それらの文字認識プログラムの認識した文字の内、最も高い評価値を示す文字を認識文字として決定したり、評価値の高い順に複数の認識文字候補を決定して、アプリケーションプログラムに渡していくという方法を用いている。

【0009】

いずれにしろ、従来では、電子化されていない文書をイメージスキャナで読み取って、それを文字認識し、誤読文字については手入力で修正していくことで文書を電子化していく場合、ユーザが電子化前の文書と認識結果とを比較することで誤読文字を検出していくという方法を探っている。

【0010】

【発明が解決しようとする課題】

しかしながら、このような従来技術に従っていると、ユーザは、電子化対象の文書と文字認識装置により認識されたその文書の認識結果とを、1つ1つ目で追いながら比較することで誤読文字を検出しなければならず、多大な負荷を強いられるという問題点があった。

【0011】

本発明はかかる事情に鑑みてなされたものであって、イメージスキャナの読み取る文書画像を処理対象として、その文書画像の持つ文字を認識する構成を探るときにとって、ユーザが認識結果の誤読文字を効率的にチェックできるようにする新たな文字認識装置及びその方法の提供と、その文字認識装置の実現に用いられるプログラムが格納される新たなプログラム記録媒体の提供とを目的とするものである。

【0012】

【課題を解決するための手段】

図1に本発明の原理構成を図示する。

【0013】

図中、1は本発明を具備する文字認識装置、2は文字認識装置1の備えるイメージスキャナ、3は文字認識装置1の備えるディスプレイである。

【0014】

本発明の文字認識装置1は、入力手段10と、入力ファイル11と、第1の認識手段12と、第1の認識用辞書13と、第1の認識結果ファイル14と、第2の認識手段15と、第2の認識用辞書16と、第2の認識結果ファイル17と、抽出手段18と、出力ファイル19と、出力手段20とを備える。

【0015】

この入力手段10は、イメージスキャナ2の読み取る文書画像を入力する。入力ファイル11は、入力手段10により入力された文書画像（例えば2値化されている）を格納する。

【0016】

第1の認識手段12は、規定の認識手法に従って、入力手段10の入力した文書画像の持つ文字を認識する。第1の認識用辞書13は、第1の認識手段12の認識処理に必要となる辞書データを管理する。第1の認識結果ファイル14は、第1の認識手段12の認識結果を格納する。

【0017】

第2の認識手段15は、第1の認識手段12の用いる認識手法とは別の認識手法に従って、入力手段10の入力した文書画像の持つ文字を認識する。第2の認識用辞書16は、第2の認識手段15の認識処理に必要となる辞書データを管理する。第2の認識結果ファイル17は、第2の認識手段15の認識結果を格納する。

【0018】

抽出手段18は、第1の認識手段12の認識結果と第2の認識手段15の認識結果との不一致個所を抽出する。

【0019】

出力ファイル19は、抽出手段18の抽出した不一致個所を記録しつつ、第1及び第2の認識手段12, 15の認識結果を格納する。出力手段20は、抽出手段18の抽出した不一致個所を明示しつつ、入力手段10の入力した文書画像の持つ文字の認識結果を出力する。

【0020】

ここで、本発明の文字認識装置1の持つ機能は具体的にはプログラムで実現されるものであり、このプログラムは、フロッピィディスクなどに格納されたり、サーバなどのディスクなどに格納され、それから文字認識装置1にインストールされてメモリ上で動作することで、本発明を実現することになる。

【0021】

このように構成される本発明の文字認識装置1では、第1の認識手段12は、規定の認識手法に従って、入力手段10の入力した文書画像の持つ文字を認識し、一方、第2の認識手段15は、第1の認識手段12の用いる認識手法とは別の認識手法に従って、入力手段10の入力した文書画像の持つ文字を認識する。

【0022】

この認識結果を受けて、抽出手段18は、第1の認識手段12の認識結果と第2の認識手段15の認識結果との不一致個所を抽出し、この抽出結果を受けて、出力手段20は、抽出手段18の抽出した不一致個所を明示しつつ、入力手段10の入力した文書画像の持つ文字の認識結果を出力する。

【0023】

この構成を探るときに、出力手段20は、文書画像と認識結果とを対比させつつその認識結果をディスプレイ画面に表示することで出力するとともに、その表示にあたって、認識結果の表示領域にカーソルを表示するときに、そのカーソル位置に対応付けられる文書画像の個所を明示する形態でその文書画像を表示することがある。

【0024】

また、この構成を探るときに、出力手段20は、認識結果が一致するものの、その認識の信頼性が低い個所については、その旨が分かる形態で出力することがある。

【0025】

このように、本発明の文字認識装置1では、イメージスキャナの読み取る文書画像を処理対象として、その文書画像の持つ文字を認識する構成を探るときにあって、別々の認識手法に従う複数の認識機能を用意し、それらの認識機能の認識結果の不一致個所を抽出して、その不一致個所を明示しつつ、文字の認識結果を

出力するという構成を探ることから、ユーザは認識結果の誤読文字がどれであるのかを一目で分かるようになることで、誤読文字を効率的にチェックできるようになる。

【0026】

【発明の実施の形態】

以下、実施の形態に従って本発明を詳細に説明する。

【0027】

図2に、本発明の文字認識装置1の一実施例を図示する。ここで、図1で説明したものと同じものについては同一の記号で示してある。

【0028】

図中、3aはディスプレイやマウスなどの入力装置を備える端末であって、ユーザとの対話手段となるもの、10aは入力プログラムであって、図1の入力手段10に相当するもの、12aは第1の認識プログラムであって、図1の第1の認識手段12に相当するもの、15aは第2の認識プログラムであって、図1の第2の認識手段15に相当するもの、18aは抽出プログラムであって、図1の抽出手段18に相当するもの、20aは出力プログラムであって、図1の出力手段20に相当するものである。

【0029】

ここで、入力プログラム10aや第1の認識プログラム12aや第2の認識プログラム15aや抽出プログラム18aや出力プログラム20aは、フロッピィディスクや回線などを介してインストールされることになる。

【0030】

このように構成される本発明の文字認識装置1は、例えば活字文字で印刷される文書を処理対象として、その文書をイメージスキャナ2で読み取り、第1の認識プログラム12aと第2の認識プログラム15aとを使って文字認識を行い、その2つの認識プログラム12a, 15aが不一致と認識した個所を明示表示しつつ、認識結果を端末3aのディスプレイ画面に表示することで、電子化されていない文書の電子化を効率的に実行できるようにする処理を行う。

【0031】

この処理を実現するために用意される第1の認識プログラム12aと第2の認識プログラム15aとは、別々の認識アルゴリズムに従って文字認識を実行したり、認識アルゴリズムそのものは同一のものに従うものの、実装レベルでの具体的な手順などに違いがある別々の方法に従って文字認識を実行するというように、別々の認識手法に従って文字認識を実行するように処理する。

【0032】

そして、これに合わせて、第1の認識プログラム12aの参照する第1の認識用辞書13と、第2の認識プログラム15aの参照する第2の認識用辞書16とが用意されることになるが、第1の認識プログラム12aの用いる認識手法と第2の認識プログラム15aの用いる認識手法とによっては、第1の認識用辞書13と第2の認識用辞書16とが共通のものとして用意されることもある。

【0033】

活字文字の認識アルゴリズムとしては、特徴抽出として、拡張セル特徴法や加重方向ヒストグラム法や外郭方向寄与度特徴法といったものが知られており、識別処理として、シティブロック距離やユークリッド距離や線形識別関数や部分空間法やKNN法やペイズ識別法などが知られている。

【0034】

これから、本発明の文字認識装置1が活字で印刷される文書を処理対象とする場合には、第1の認識プログラム12aや第2の認識プログラム15aは、例えば、これらの認識アルゴリズムの中から選択される別々の認識アルゴリズムを使って文字認識を実行することになる。

【0035】

図3に、第1及び第2の認識プログラム12a, 15aの実行する処理フローの一実施例、図4及び図5に、抽出プログラム18aの実行する処理フローの一実施例、図6に、出力プログラム20aの実行する処理フローの一実施例を図示する。

【0036】

次に、これらの処理フローに従って、本発明の文字認識装置1の実行する処理について詳細に説明する。以下、説明の便宜上、入力プログラム10aの入力す

る処理対象の文書画像が横書きであることを想定する。

【0037】

最初に、図3の処理フローに従って、第1及び第2の認識プログラム12a, 15aの実行する処理について説明する。

【0038】

第1及び第2の認識プログラム12a, 15aは、入力ファイル11に格納される文書画像を指定して処理要求が発行されると、図3の処理フローに示すように、先ず最初に、ステップ1で、入力ファイル11から処理対象の文書画像を読み出し、公知の切り出しアルゴリズムに従って、行を単位にして、その読み出した文書画像の中から文字領域のイメージデータを切り出す。

【0039】

すなわち、図7(a)に示すように、横書きの文書画像に合わせて、行を単位にして、文書画像の中から文字領域のイメージデータを切り出すのである。ここで、処理対象の文書画像が縦書きである場合には、列を単位にして、文書画像の中から文字領域のイメージデータを切り出すことになる。なお、文書画像が横書きであるのか縦書きであるのかについては、入力プログラム10aが文書画像を読み込むときに、ユーザやアプリケーションプログラムなどから取得することになる。

【0040】

続いて、ステップ2で、公知の切り出しアルゴリズムに従って、その行単位に切り出した文字領域のイメージデータから、文書画像に記載される各文字の文字領域を切り出して、それらのイメージデータ上の位置を検出する。

【0041】

すなわち、図7(b)に示すように、行単位に切り出した文字領域のイメージデータから、四角形で定義される各文字の文字領域を切り出し、例えば、その各文字領域のイメージデータ上の左上位置座標(x_{j1}, y_{j1})と右下位置座標(x_{j2}, y_{j2})とを検出することで、各文字領域のイメージデータ上の位置を検出するのである。

【0042】

続いて、ステップ3で、規定の認識手法を使って、そのようにして切り出した各文字領域の文字コードを認識する。このとき、第1位の評価値（類似度）を持つ文字を認識結果（認識文字）とするが、その認識文字に対応付けて、高い評価値を持つ規定の数の文字を認識候補文字として登録する。ここで、上述したように、第1の認識プログラム12aと第2の認識プログラム15aとは、別々の認識手法を使って文字認識を実行することになる。

【0043】

続いて、ステップ4で、認識結果となる認識文字毎に、行位置／文字領域のイメージデータ位置／文字コード（認識文字とそれにリンクする認識候補文字の文字コード）／評価値（認識文字とそれにリンクする認識候補文字の評価値）を対データとする認識情報を作成して、それを認識結果ファイル14, 17に格納して処理を終了する。

【0044】

このようにして、第1の認識結果ファイル14には、第1の認識プログラム12aにより作成された認識情報（各認識文字の行位置／文字領域のイメージデータ位置／文字コード／評価値）が格納され、第2の認識結果ファイル17には、第2の認識プログラム15aにより作成された認識情報（各認識文字の行位置／文字領域のイメージデータ位置／文字コード／評価値）が格納されることになる。

【0045】

次に、図4及び図5の処理フローに従って、抽出プログラム18aの実行する処理について説明する。

【0046】

抽出プログラム18aは、第1及び第2の認識結果ファイル14, 17に認識情報が格納されることで起動されると、図4及び図5の処理フローに示すように、先ず最初に、ステップ1で、行番号を示す変数iに初期値“1”をセットする。

【0047】

続いて、ステップ2で、変数iが行番号の最大値imax(図7(a)に示すもの

) を超えたのか否かを判断して、超えていないことを判断するときには、ステップ3に進んで、第1の認識結果ファイル14から、第1の認識プログラム12aの作成したi行に関する認識情報を読み込むとともに、第2の認識結果ファイル17から、第2の認識プログラム15aの作成したi行に関する認識情報を読み込む。なお、この時点では、認識候補文字に関する情報については読み込む必要はない。

【0048】

続いて、ステップ4で、第1の認識結果ファイル14から読み込んだ各文字領域の左上位置座標(x_{j1}, y_{j1})の持つ座標 x_{j1} ($j = 1 \sim j_{max}$)のリストと、第2の認識結果ファイル17から読み込んだ各文字領域の左上位置座標(x'_{j1}, y'_{j1})の持つ座標 x'_{j1} ($j = 1 \sim j'_{max}$)のリストとから、「 $x_{j1} = x'_{j1}$ 」とならないものがあるのか否かを判断する。

【0049】

この判断処理に従って、全ての座標 x_{j1}, x'_{j1} について、「 $x_{j1} = x'_{j1}$ 」が成立することを判断するとき、すなわち、第1の認識プログラム12aの認識したi行の認識文字の数と、第2の認識プログラム15aの認識したi行の認識文字の数とが等しいことを判断するときには、ステップ5に進んで、i行の各認識文字毎に、第1の認識プログラム12aの認識した認識文字と、第2の認識プログラム15aの認識した認識文字とが一致するのか否かを判断することで、不一致となる認識文字の文字位置を検出するとともに、その不一致となる2つの認識文字の持つ評価値を比較して、評価値の高い方の認識文字を表示対象として設定する。

【0050】

すなわち、図8に示すように、第1の認識プログラム12aの認識した認識文字と、第2の認識プログラム15aの認識した認識文字とが不一致となる場合には、その2つの認識文字の内の評価値の高い方の認識文字を表示対象として設定するのである。

【0051】

続いて、ステップ6に進んで、次の行の処理に進むべく、変数iの値を1つイ

ンクリメントしてから、ステップ2に戻る。

【0052】

一方、ステップ4で、全ての座標 x_{j1}, x_{j1}' について、「 $x_{j1} = x_{j1}'$ 」が成立しないことを判断するとき、すなわち、第1の認識プログラム12aの認識したi行の認識文字の数と、第2の認識プログラム15aの認識したi行の認識文字の数とが等しくないことを判断（但し、土の関係で偶然等しくなることもある）するときは、ステップ7（図5の処理フロー）に進んで、等しくないイメージデータ位置を挟む2つのイメージデータ位置（このイメージデータ位置では、「 $x_{j1} = x_{j1}'$ 」が成立する）を特定する。

【0053】

すなわち、漢字の「化」を、片仮名の「イ」と片仮名の「ヒ」の2文字であると誤認識したり、片仮名の「ノ」と片仮名の「レ」との2文字を、片仮名の「ル」の1文字と誤認識したりすることで、第1の認識プログラム12aの認識したi行の認識文字の数と、第2の認識プログラム15aの認識したi行の認識文字の数とが等しくならないことが起こる。

【0054】

このような場合には、図9に示すように、「 $x_{j1} = x_{j1}'$ 」が成立しないイメージデータ位置（図中の①）があるので、そのイメージデータ位置を挟む2つのイメージデータ位置（図中の②と③）を特定するのである。この2つのイメージデータ位置では、「 $x_{j1} = x_{j1}'$ 」が成立している。

【0055】

続いて、ステップ8で、その特定したイメージデータ位置の間にある認識文字を不一致と判断するとともに、文字数の多い方の認識文字を表示対象として設定する。例えば、この不一致個所の認識文字として、第1の認識プログラム12aが「ノ」と「レ」と認識し、第2の認識プログラム15aが「ル」と認識したときには、「ノ」と「レ」とを表示対象として設定するのである。

【0056】

ここで、文字数の多い方の認識文字を表示対象とするのではなくて、文字数の少ない方の認識文字を表示対象として設定する方法を探ってもよい。例えば、こ

の不一致個所の認識文字として、第1の認識プログラム12aが「ノ」と「レ」と認識し、第2の認識プログラム15aが「ル」と認識したときには、「ル」を表示対象として設定するという方法を探ってもよい。

【0057】

また、 i 行の持つ各文字領域の左上位置座標の持つ座標 x_{j1}, x_{j1}' のリストから1文字の大きさを推定し、その推定した大きさから不一致個所の認識文字の文字数を判断して、それに従って、第1の認識プログラム12aの認識した不一致個所の認識文字か、第2の認識プログラム15aの認識した不一致個所の認識文字かを選択して、それを表示対象として設定するという方法を探ってもよい。この方法を用いると、処理は若干複雑になるものの、実際のものである可能性の高い認識文字を表示対象として設定できるようになる。

【0058】

いずれにしろ、このステップ8では、ステップ7で検出した不一致個所の認識文字として、何らかの判断基準に従って、第1の認識プログラム12aの認識した不一致個所の認識文字か、第2の認識プログラム15aの認識した不一致個所の認識文字かを選択して、それを表示対象として設定するように処理するのである。

【0059】

続いて、ステップ9で、等しいイメージデータ位置を持つ認識文字の中から、不一致となる認識文字の文字位置を検出するとともに、その不一致となる2つの認識文字の持つ評価値を比較して、評価値の高い方の認識文字を表示対象として設定する。

【0060】

すなわち、ステップ7で検出した不一致個所以外の個所については、同一のイメージデータ位置に対応付けられる認識文字が存在するので、その認識文字同士が不一致となるのか否かを検出して、不一致となる認識文字については、評価値の高い方の認識文字を表示対象として設定するのである。

【0061】

続いて、ステップ10に進んで、次の行の処理に進むべく、変数*i*の値を1つ

インクリメントしてから、ステップ2に戻る。

【0062】

このようにして、ステップ2ないしステップ10の処理を繰り返していくことで、変数 i が行番号の最大値 i_{max} (図7 (a) に示すもの) を超えたことを判断すると、ステップ11 (図4の処理フロー) に進んで、以上の処理により特定した、一致した認識文字 (表示対象となる) と、抽出した不一致個所と、不一致個所での表示対象となる認識文字と、表示対象となる認識文字のイメージデータ位置とを出力ファイル19に格納するとともに、第1及び第2の認識結果ファイル14, 17から、不一致個所における認識文字及び認識候補文字とそれが持つ評価値とを読み込んで、それを出力ファイル19に格納して、処理を終了する。

【0063】

次に、図6の処理フローに従って、出力プログラム20aの実行する処理について説明する。

【0064】

出力プログラム20aは、抽出プログラム18aの処理に従って、出力ファイル19に認識結果の情報が格納されると起動されると、図6の処理フローに示すように、先ず最初に、ステップ1で、入力ファイル11から、表示対象となる文書画像のイメージデータを読み込み、続くステップ2で、その読み込んだイメージデータを、端末3aのディスプレイ画面に開設される左側ウィンドウに表示する。

【0065】

続いて、ステップ3で、出力ファイル19から、表示対象となる文書画像の認識結果情報を読み込み、続くステップ4で、その読み込んだ認識結果情報の指定する一致個所での認識文字と、その読み込んだ認識結果情報の指定する不一致個所での表示対象となる認識文字とを異なる表示形態を使って、端末3aのディスプレイ画面に開設される右側ウィンドウに表示するとともに、その右側ウィンドウにカーソルを表示する。

【0066】

続いて、ステップ5で、ステップ3で読み込んだ認識結果情報に従って、カーソル

ソルの指す認識文字のイメージデータ位置を取得して、そのイメージデータ位置が指定する左側ウィンドウの対応する位置が分かるようにと明示表示して、処理を終了する。

【0067】

図10に、出力プログラム20aが表示するディスプレイ画面の一実施例を図示する

この図では、①に記載される「機」という文字が、第1の認識プログラム12aと第2の認識プログラム15aとで認識結果が異なり、図4の処理フローのステップ5の処理に従って、評価値の高い方の認識文字である「機」が表示対象として設定されることで表示されることになることを想定している。

【0068】

また、②に記載される「化」という文字が、第1の認識プログラム12aと第2の認識プログラム15aとで認識結果が異なり、図5の処理フローのステップ8の処理に従って、文字数の多い方の認識文字である「イ」と「ヒ」とが表示対象として設定されることで表示されることになることを想定している。

【0069】

また、③に記載される「文」という文字に「▲」で示すカーソルを表示するときには、左ウィンドウに表示する文書画像のイメージデータ中の対応する位置に、カーソルの指すイメージデータ上の位置であることを示す「△」を表示することを想定している。

【0070】

この図では表示できないが、出力プログラム20aは、第1の認識プログラム12aと第2の認識プログラム15aとで認識結果が同一となった認識文字と、①や②に表示する不一致個所の認識文字とを、その旨が分かるようにと、色などを変えた異なる表示形態で表示している。このとき、更に、①に表示する認識文字と②に表示する認識文字とが異なる種類の誤読文字であることを表すために、その表示形態を変えるようにする構成を探ることも可能である。

【0071】

すなわち、①に表示する認識文字は、第1の認識プログラム12aと第2の認

識プログラム15aとで行内の文字数については一致した誤読文字であり、②に示す認識文字は、第1の認識プログラム12aと第2の認識プログラム15aとで行内の文字数が異なるものとなった認識文字であるので、それが分かるようにと、①に表示する認識文字と②に表示する認識文字との表示形態を変えるようにする構成を探ることも可能である。

【0072】

そして、このような表示処理を実行するときにとって、出力プログラム20aは、ユーザが認識文字の不一致個所である①や②に表示する認識文字をクリックするときには、出力ファイル19から読み込んだ、不一致個所における認識文字及び認識候補文字とそれが持つ評価値とから、評価値の高い順にそれらの認識文字及び認識候補文字をソートして、そのソート順に認識文字（認識候補文字）を順次表示していくように処理することになる。

【0073】

この表示による選択処理でも本来の文字を表示できない場合には、出力プログラム20aは、図示しないエディタと連携して、ユーザからの文字入力に応答して、誤読文字を本来の文字に置き換えていくように処理する。

【0074】

図10に示した実施例では、出力プログラム20aは、第1の認識プログラム12aと第2の認識プログラム15aとで認識結果が異なる個所については、いずれか一方の認識した認識文字を表示する構成を探っているが、図11に示すように、認識文字を表示する代わりに、その旨が分かる記号（空白でもよい）を表示することによってもよい。

【0075】

以上に説明した実施例では説明しなかったが、第1の認識プログラム12aと第2の認識プログラム15aとで認識結果が一致するものの、いずれか一方又は双方での評価値が低い場合には、その認識結果の認識文字の信頼性が低いことを判断して、高い信頼性を示す認識文字とは異なる表示形態で表示することで、ユーザに注意を促すという構成を探ることも可能である。

【0076】

図示実施例に従って本発明を説明したが、本発明はこれに限定されるものではない。例えば、実施例では、出力装置としてディスプレイを用いる構成を採ったが、本発明は、出力装置としてプリンタなどを用いる場合にも、そのまま適用できる。プリンタに出力する場合には、書体などを変えることで不一致の認識文字であることが分かるように出力することになる。

【0077】

また、実施例では、認識文字のイメージデータ位置を使って、第1の認識プログラム12aと第2の認識プログラム15aとの認識した文字が一致するのか否かを判断するという構成を採ったが、認識文字列の並びをマッチングすることなどにより判断するというように、認識文字のイメージデータ位置を使わない別の方法を用いることも可能である。

【0078】

また、実施例では、第1の認識プログラム12aと第2の認識プログラム15aという2つの認識プログラムを用いる構成を採ったが、本発明は、3つ以上の認識プログラムを用いる場合にも、そのまま適用できる。

【0079】

また、本発明は、文字種別に関係なく適用可能であり、例えば、活字文字にその適用が限られるものではなく、手書き文字に対してもそのまま適用できる。

【0080】

【発明の効果】

以上説明したように、本発明の文字認識装置では、イメージスキャナの読み取る文書画像を処理対象として、その文書画像の持つ文字を認識する構成を探るときにあって、別々の認識手法に従う複数の認識機能を用意し、それらの認識機能の認識結果の不一致個所を抽出して、その不一致個所を明示しつつ、文字の認識結果を出力するという構成を探ることから、ユーザは認識結果の誤読文字がどれであるのかを一目で分かるようになることで、誤読文字を効率的にチェックできるようになる。

【図面の簡単な説明】

【図1】

本発明の原理構成図である。

【図2】

本発明の一実施例である。

【図3】

認識プログラムの処理フローである。

【図4】

抽出プログラムの処理フローである。

【図5】

抽出プログラムの処理フローである。

【図6】

出力プログラムの処理フローである。

【図7】

認識プログラムの処理の説明図である。

【図8】

抽出プログラムの処理の説明図である。

【図9】

抽出プログラムの処理の説明図である。

【図10】

出力プログラムの処理の説明図である。

【図11】

出力プログラムの処理の説明図である。

【符号の説明】

- 1 文字認識装置
- 2 イメージスキャナ
- 3 ディスプレイ
- 10 入力手段
- 11 入力ファイル
- 12 第1の認識手段
- 13 第1の認識用辞書

1 4 第1の認識結果ファイル

1 5 第2の認識手段

1 6 第2の認識用辞書

1 7 第2の認識結果ファイル

1 8 抽出手段

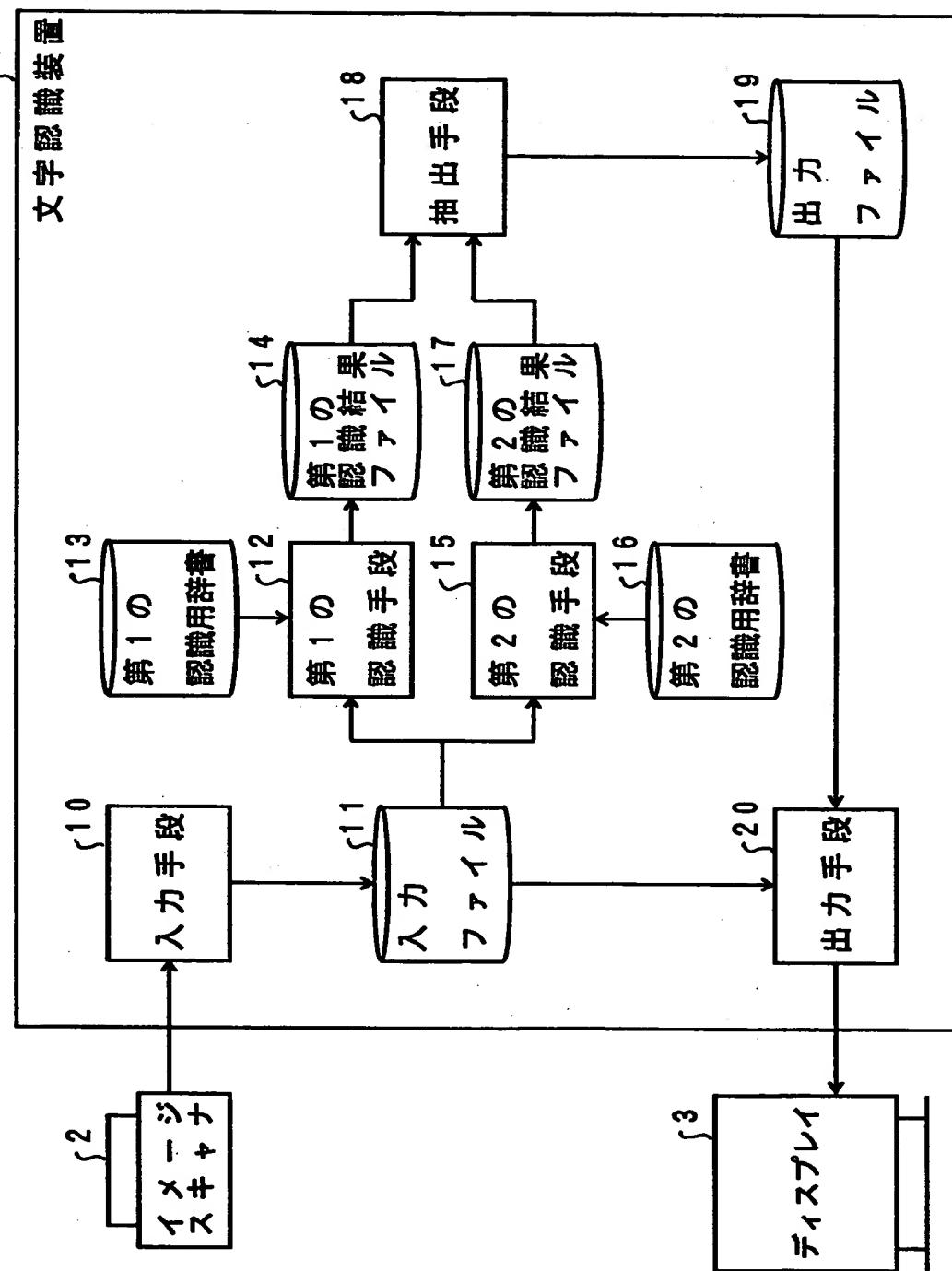
1 9 出力ファイル

2 0 出力手段

【書類名】 図面

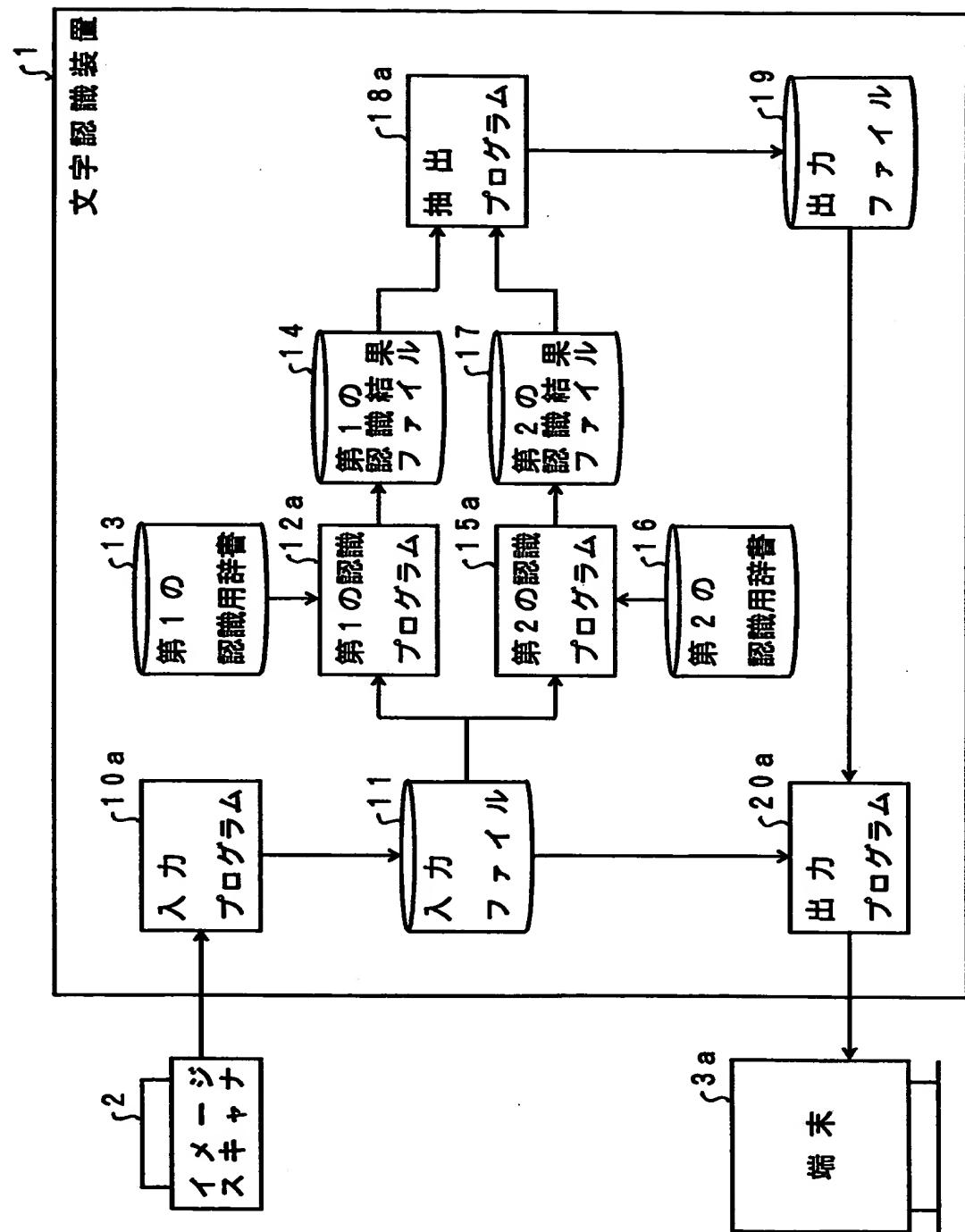
【図1】

本発明の原理構成図



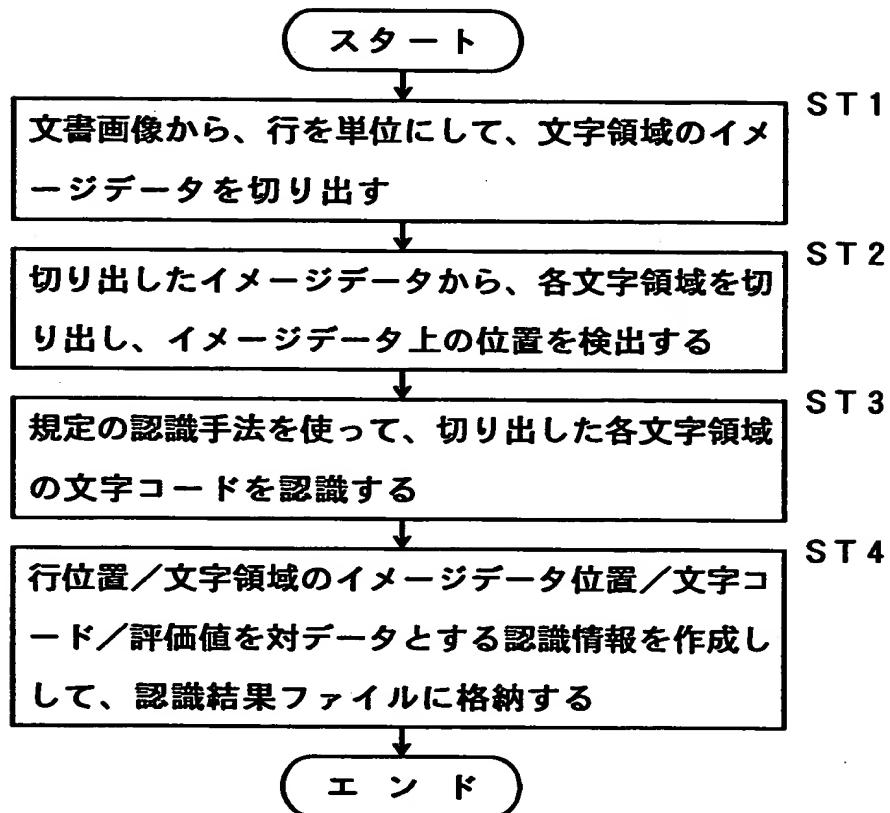
【図2】

本発明の一実施例



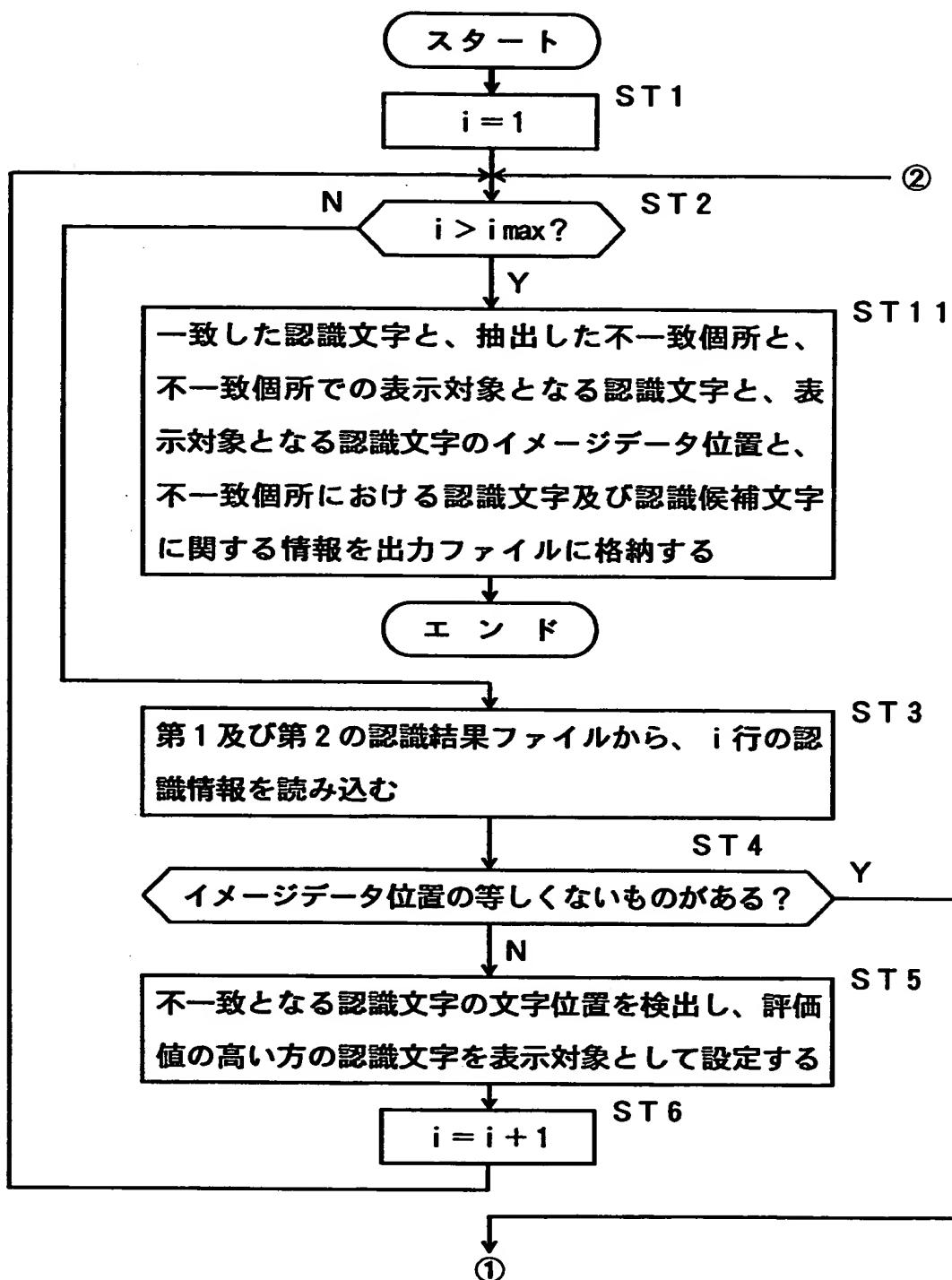
【図3】

認識プログラムの処理フロー



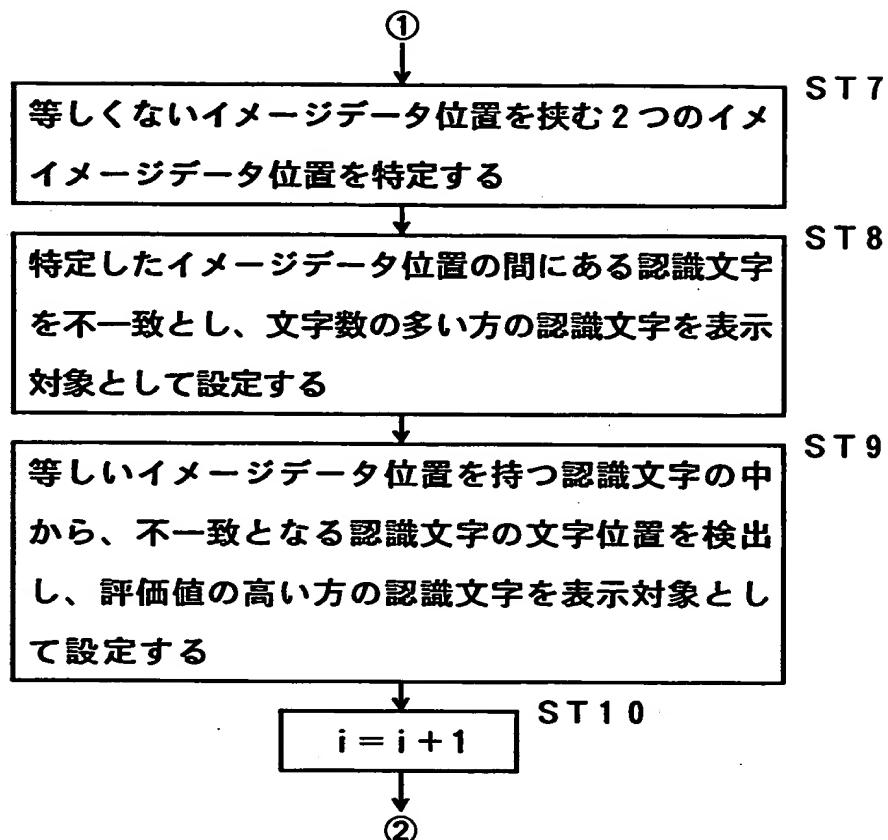
【図4】

抽出プログラムの処理フロー



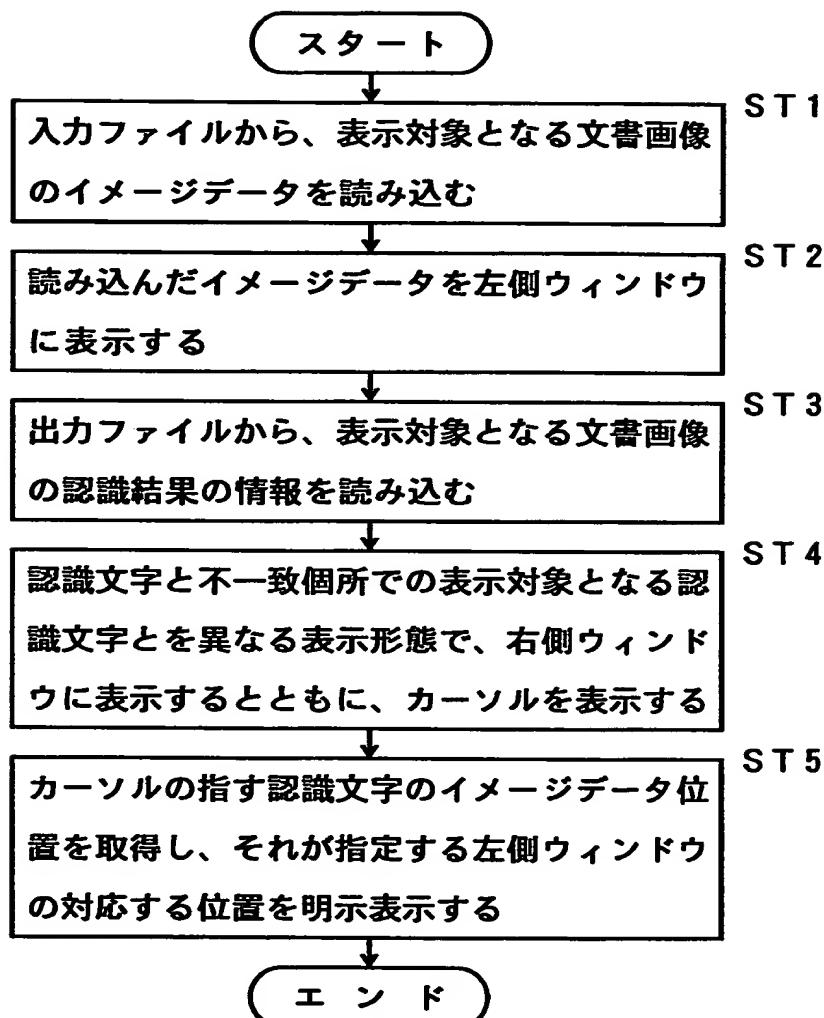
【図5】

抽出プログラムの処理フロー



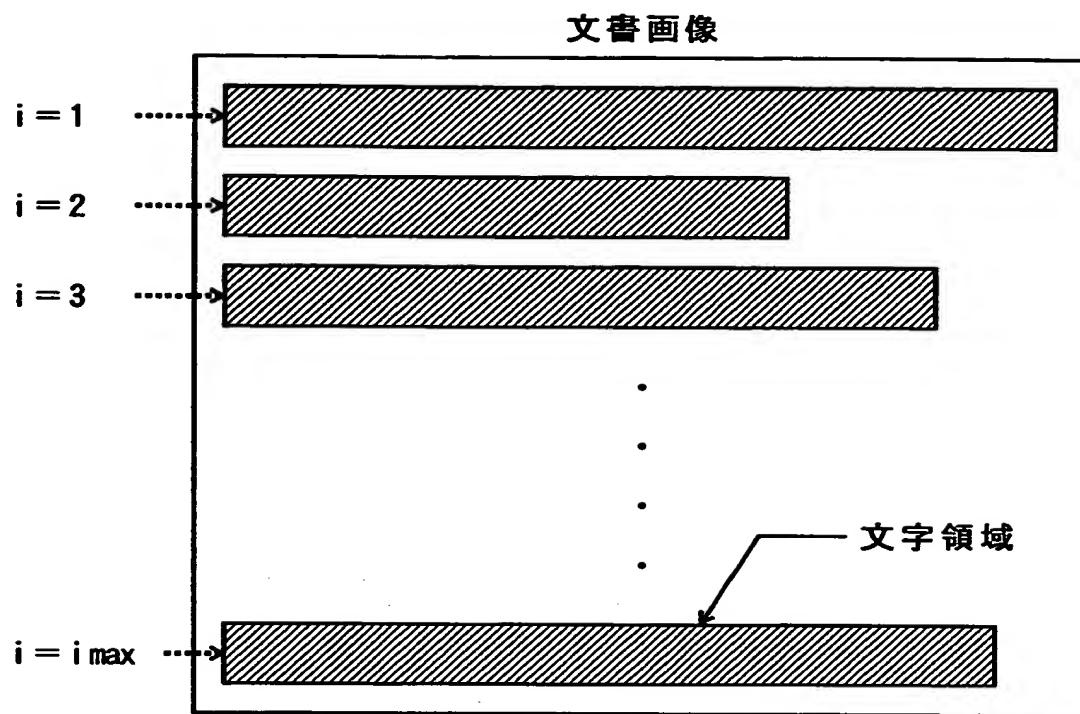
【図6】

出力プログラムの処理フロー

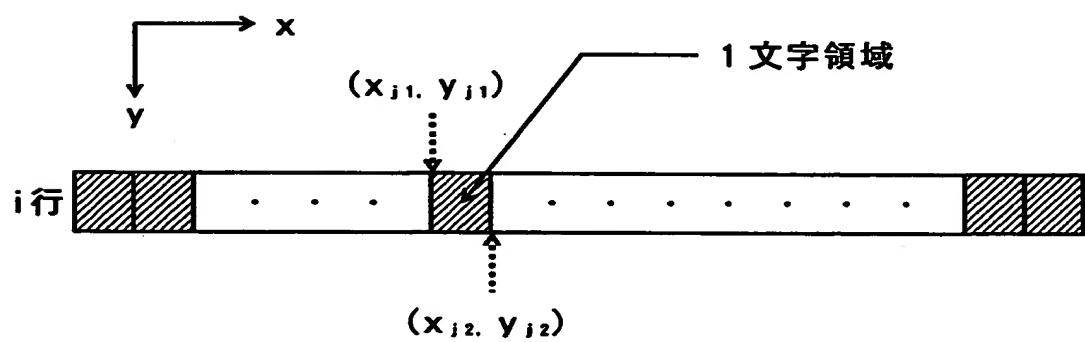


【図7】

認識プログラムの処理の説明図



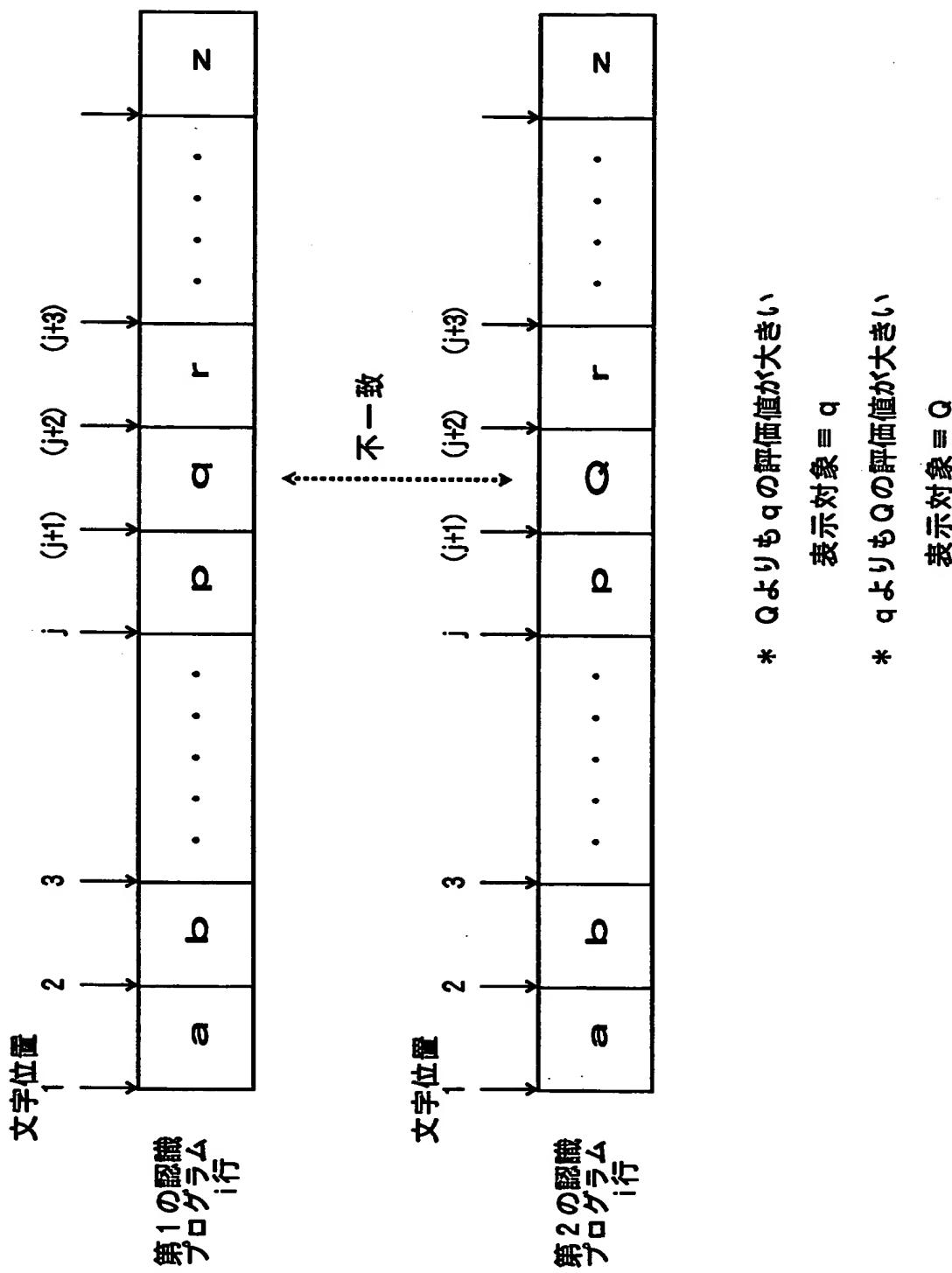
(a)



(b)

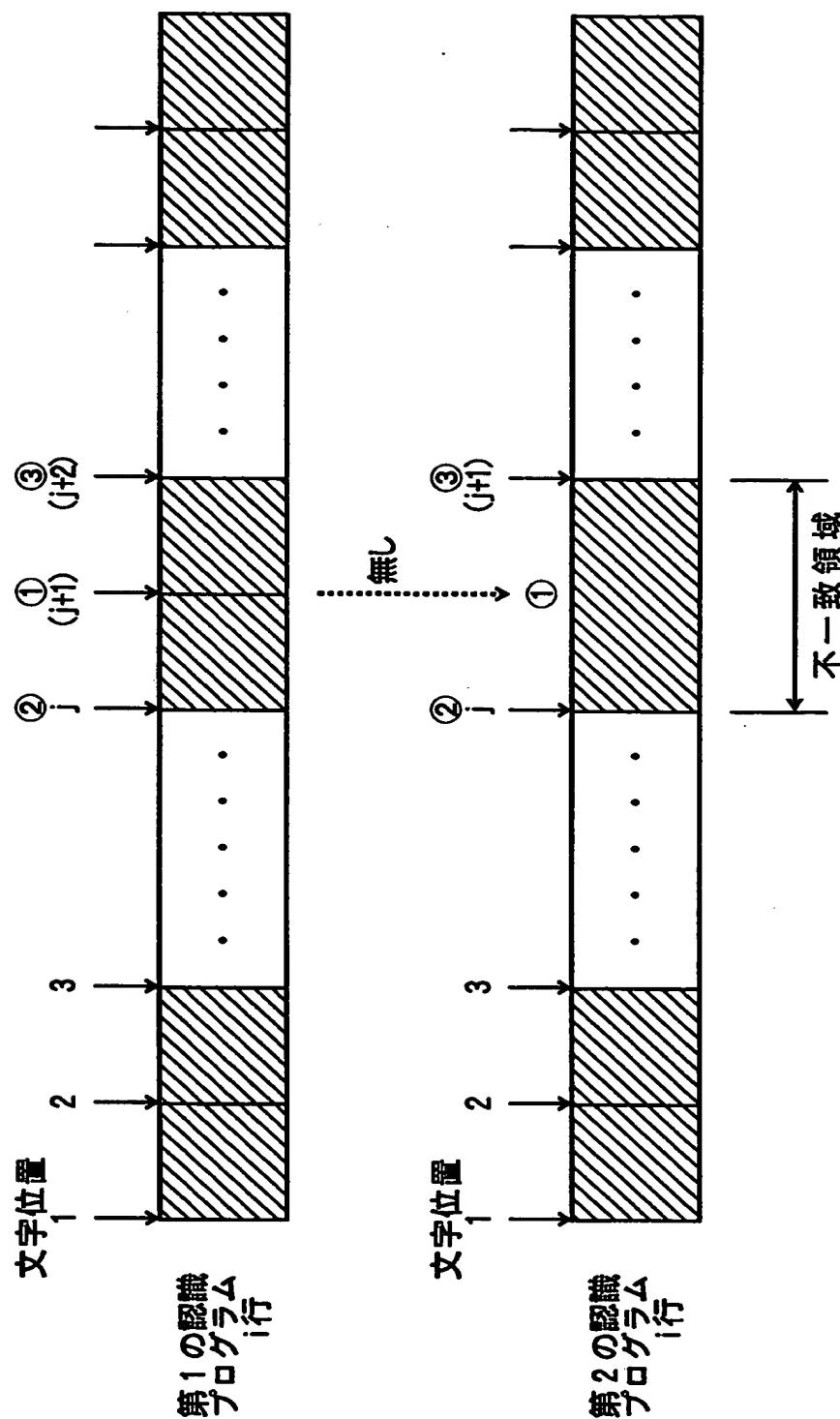
【図8】

抽出プログラムの処理の説明図



【図9】

抽出プログラムの処理の説明図



【図10】

圧力プログラムの処理の説明図

イメージデータ	認識結果
□ □ □	<p>パターン認識は文字を読み音を聞き分けると いった人間の能力を機械で実現するための要求 から出発し、これまでに多くの研究が行われて きた。とりわけ文字認識は最も歴史が古く、か つ実用化が進んだ分野である。</p> <p>文字は身近な存在であり、記録性に優れ、再 現性が良く、人間の直感を生かしやすく、1文 字に一つの概念が対応するなど、パターン認識 研究の素材としては格好の性質を備えている。</p> <p>また文字を読む技術は、コンピュータへのデ ータ入力をキーボードを経ずに直接行う省力装</p> <p>パターン認識は文字を読み音を聞き分けると いった人間の能力を機械で実現するための要求 から出発し、これまでに多くの研究が行われて きた。とりわけ文字認識は最も歴史が古く、か つ実用化が進んだ分野である。</p> <p>文字は身近な存在であり、記録性に優れ、再 現性が良く、人間の直感を生かしやすく、1文 字に一つの概念が対応するなど、パターン認識 研究の素材としては格好の性質を備えている。</p> <p>また文字を読む技術は、コンピュータへのデ ータ入力をキーボードを経ずに直接行う省力装</p>

【図11】

比力プログラムの処理の説明図

イメージデータ	認識結果
<p>パターン認識は文字を読み音を聞き分けると いった人間の能力を機械で実現するための要求 から出発し、これまでに多くの研究が行われて きた。とりわけ文字認識は最も歴史が古く、か つ実用化が進んだ分野である。</p> <p>文字は身近な存在であり、記録性に優れ、再 現性が良く、人間の直感を生かしやすく、1文 字に一つの概念が対応するなど、パターン認識 研究の素材としては恰好の性質を備えている。</p> <p>また文字を読み技術は、コンピュータへのデ ータ入力をキーボードを経ずに直接行う省力装</p>	<p>パターン認識は文字を読み音を聞き分けると いった人間の能力を^①機械で実現するための要求 から出発し、これまでに多くの研究が行われて きた。とりわけ文字認識は最も歴史が古く、か つ実用^②が進んだ分野である。</p> <p>文字は身近な存在であり、記録性に優れ、再 現性が良く、人間の直感を生かしやすく、1文 字に一つの概念が対応するなど、パターン認識 研究の素材としては格好の性質を備えている。</p> <p>また文字を読み技術は、コンピュータへのデ ータ入力をキーボードを経ず^③に直接行う省力装</p>

【書類名】 要約書

【要約】

【課題】 本発明は、イメージスキャナの読み取る文書画像を処理対象として、その文書画像の持つ文字を認識する文字認識装置に関し、ユーザが認識結果の誤読文字を効率的にチェックできるようにすることを目的とする。

【解決手段】 規定の認識手法に従って、処理対象の文書画像の持つ文字を認識する第1の認識手段12と、第1の認識手段12の用いる認識手法とは別の認識手法に従って、処理対象の文書画像の持つ文字を認識する第2の認識手段15と、第1の認識手段12の認識結果と第2の認識手段15の認識結果との不一致個所を抽出する抽出手段18と、抽出手段18の抽出する不一致個所を明示しつつ、処理対象の文書画像の持つ文字の認識結果を出力する出力手段20とを備えるように構成する。

【選択図】 図1

出願人履歴情報

識別番号 [000005223]

1. 変更年月日 1996年 3月26日

[変更理由] 住所変更

住 所 神奈川県川崎市中原区上小田中4丁目1番1号
氏 名 富士通株式会社

出願人履歴情報

識別番号 [599131675]

1. 変更年月日 1999年 9月16日

[変更理由] 新規登録

住 所 フィリピン国、セブ市 6000、バニラッド、パセオアナベ
ル マリア ルイサ エステート パーク、18

氏 名 ジャンプ セブ インコーポレーション